

ARI Research Note 99-24

**Optimizing the Long-Term Retention of Skills:
Structural and Analytic Approaches to Skill Maintenance
Annual Report, 1991-1992**

Alice F. Healy, K. Anders Ericsson, and Lyle E. Bourne, Jr.
University of Colorado

**Research and Advanced Concepts Office
Michael Drillings, Chief**

April 1999



**U.S. Army Research Institute
for the Behavioral and Social Sciences**

Approved for public release; distribution is unlimited.

19990422 032

**U.S. Army Research Institute
for the Behavioral and Social Sciences**

A Directorate of the U.S. Total Army Personnel Command

**EDGAR M. JOHNSON
Director**

Research accomplished under contract
for the Department of the Army

University of Colorado

Technical review by

Tony Papa

NOTICES

DISTRIBUTION: This Research Note has been cleared for release to the Defense Technical Information Center (DTIC) to comply with regulatory requirements. It has been given no primary distribution other than to DTIC and will be available only through DTIC or the National Technical Information Service (NTIS).

FINAL DISPOSITION: This Research Note may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The views, opinions, and findings in this Research Note are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY (Leave Blank)		2. REPORT DATE April 1999		3. REPORT TYPE AND DATES COVERED Annual 5/1/91 - 4/30/92	
4. TITLE AND SUBTITLE Optimizing the Long-Term Retention of Skills: Structural and Analytic Approaches to Skill Maintenance				5. FUNDING NUMBERS PE 0601102A 20161102B74F TA 2901 WU C07 Contract No. MDA903-90-K-0066	
6. AUTHOR(S) Alice Healy, K. Anders Ericsson, & Lyle Bourne, Jr. (University of Colorado)					
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Colorado Campus Box B-19 Boulder CO 80309 (303)492-5032				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U. S. Army Research Institute ATTN: PERI-BR 5001 Eisenhower Avenue Alexandria, VA 22333-5600				10. SPONSORING/MONITORING AGENCY REPORT NUMBER Research Note 99-24	
11. SUPPLEMENTARY NOTES COR: Dr. George Lawton					
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for Public Release; Distribution is Unlimited.				12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words): This research program seeks to identify the characteristics of knowledge and skill which are most resistant to decay due to disuse. The general goal is to elucidate principles which will specify those aspects of a complex skill that resist decay over periods of disuse and how they are distinguishable from more fragile components. The research program can be divided into two complementary parts. The first part is concerned with describing the structure of existing skills. The second part is concerned with experimental analysis of factors influencing and improving retention of skill components. Our work encompassed a large number of different studies on a wide range of tasks, including tank gunner skills, Morse code reception, color naming, instrument panel scanning, mental calculation, memory for instances of categories, target detection, data entry, components of memory for lists, components of memory for schedules, and vocabulary retention. Each of these tasks provided a test bed for our major theoretical hypothesis that the durability of memory depends critically on the extent to which learning procedures are reinstated at test.					
14. SUBJECT TERMS practice retention "permastore" knowledge skill					15. NUMBER OF PAGES
					16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited		

9000660103953

Annual Report for the period May 1, 1991 to April 30, 1992

Optimizing the Long-term Retention of Skills:

Structural and Analytic Approaches to Skill Maintenance

Principal Investigator: Alice F. Healy

Co-Investigators: K. Anders Ericsson and Lyle E. Bourne, Jr.

Postdoctoral Research Associates: Vivian Schneider and Grant Sinclair

Graduate Research Assistants: Deborah Clawson, Robert Crutcher,

Oliver Hammerle, Cheri King, William Marmie,

Danielle McNamara, and Tim Rickard

Undergraduate Research Assistants: Kris Kulka, David McCain,
and Gregory Rully

Meetings. In June Alice Healy attended a meeting of the Publications Committee of the Psychonomic Society in Austin, Texas.

In July Lyle Bourne presented a paper at the Inter-American Congress of Psychology, San Jose, Costa Rica.

In August Anders Ericsson made a full-day presentation at the summer course on "Qualitative Methods" arranged by the Nordic Summer Academy for Graduate Education, Grabbestad, Sweden. This presentation was on the topic of "Verbal data in the study of cognition: Protocol analysis of cognitive processes and analysis of diaries about cognitive activities."

Anders Ericsson also delivered invited colloquia at the University of Gotenburg in Sweden (August), at the Armstrong Laboratory of the United States Air Force in Phoenix, Arizona (September), at the Department of Psychology of the University of Colorado in Boulder (October), and at St. Lawrence University in Canton, New York (October). These colloquia were all on the topic of "Expert performance: Its structure and acquisition."

Alice Healy delivered an invited colloquium at the Institute of Cognitive Science at the University of Colorado in September. This talk was prepared as an invited address to be delivered at the American Psychological Association convention in August. Due to a serious illness, Dr. Healy was unable to deliver the address at the convention. This colloquium was on the topic of "Cognitive processes in reading text."

Lyle Bourne attended the American Psychological Association convention in August. He participated as President of Division 3 (Experimental Psychology). In October he also gave invited colloquia at the University of Central Washington and at the University of Utah entitled, "Mental arithmetic: Facts and figuring," based on research supported by this contract.

Lyle Bourne spent time during his sabbatical leave at the University of California, Berkeley (November 1-20) and at the University of California, Santa Barbara (December 1-12). On November 15 he presented a colloquium at the University of California, Berkeley entitled, "Mental arithmetic: Facts and figuring." Later, on December 5, he gave the same colloquium at the University of California, Santa Barbara.

In November, Anders Ericsson presented an invited address entitled, "Prudent practice makes perfect" at the Nutrition Summit Conference "Genetics, how much of mother nature can be overridden?" The conference was held in Newport Beach, California.

Also in November, Alice Healy, Lyle Bourne, Anders Ericsson, and Danielle McNamara attended the annual meeting of the Psychonomic Society in San Francisco. Alice Healy presented a paper by Danielle McNamara and herself entitled, "A generation advantage for multiplication skill and nonword vocabulary acquisition." Anders Ericsson presented a poster by Ralf Krampe and himself entitled, "Tapping speed and bimanual coordination in experts vs. amateur pianists: Talent or practice?" Alice Healy, Tom Cunningham (of St. Lawrence University), and Lyle Bourne were coauthors of a paper presented by Robert Till (of the University of North Dakota) entitled, "Short-term memory for item, temporal, and spatial information in young and elderly." Alice Healy, Robert Till, David Fendrich (of Widener University), and Christina Dimitry (of St. Lawrence University) were also coauthors of a paper presented by Tom Cunningham entitled, "Isolating memory processes: Expectancy and item value in recall." Alice Healy and Lyle Bourne attended a meeting of the Executive Committee of Division 3 (Experimental Psychology) of the American Psychological Association, and Alice Healy also attended meetings of the Governing Board of the Psychonomic Society, the Publications Committee of the Psychonomic Society, and the Editorial Board of the journal Memory & Cognition.

In January, Robert Crutcher delivered an invited colloquium at the Department of Psychology of the University of Illinois, Chicago. The title of his talk was, "Memory retrieval of acquired knowledge as a function of practice: A process-centered approach." Robert Crutcher gave the same talk again in February at Denver University.

Also in January, Danielle McNamara presented a poster, "Revising science texts to improve comprehension," by herself, Walter Kintsch, Eileen Kintsch, and Nancy Songer at the third annual Winter Text Conference in Jackson, Wyoming.

In February, Anders Ericsson gave an invited lecture for the Conradi Eminent Scholar Chair at Florida State University with the title, "The study of expert memory performance: Implications for the structure of memory in everyday life."

Later in February, Alice Healy and Anders Ericsson attended the annual In-Progress Review Meeting for our contract and presented a report summarizing the findings from our project during the last year. This meeting was held in Winter Park, Florida.

In March, Lyle Bourne visited the Department of Clinical and Physiological Psychology at the Medical School of the University of Tübingen in West Germany. He delivered three colloquia there. The first was entitled, "Mental arithmetic: Facts and figuring." The second was entitled, "Acquisition and long-term retention of skills." And the third was entitled, "Perceptual (cognitive) and motoric components of repetition and priming effects." Lyle Bourne

also delivered the first of these talks again at the Department of Psychology, University of Aberdeen, Scotland.

Again in April, Lyle Bourne delivered colloquia entitled, "Mental arithmetic: Facts and figuring," and "Perceptual (cognitive) and motoric components of repetition and priming effects." These talks were delivered at the German Research Center for Artificial Intelligence in Kaiserslautern, Germany.

Also in April, Vivian Schneider delivered a talk on letter detection at the University of Colorado at Denver.

Tank gunner skills. Final analyses have been performed for several measures of the long-term retention experiment in our series on tank gunner skills. This experiment tested the long-term retention of a complex skill acquired by six subjects given twelve initial training sessions. Analyses of variance were performed on accuracy measures of hits, kills, and misses for serial position-related and range-related variables. A further analysis, similar to the method of Hart, Hagman, and Bowne (ARI Research Report 1560), examined the percentage of hits, kills, and misses out of the total rounds fired. The percentage was determined for each of the ten blocks of a given session. An analysis of variance was also conducted on number of penalties assessed. Results show that the presentation serial position of a given target is a less important variable than the range (i.e., the distance away from the subject). The range variable consistently affected performance for the kills, misses, and hits measures, whereas the serial position variable did not. Main effects of both session and block were also found for each of these three measures as well as significant interactions of session and block. An interaction of session and range suggests that generally there is a reduction in the magnitude of the effect of range as the number of misses and hits decrease while the number of kills increase, with some interesting exceptions. Penalties showed significant main effects of session and block when the serial position analysis was conducted. Interestingly, an interaction of session and range emerged in the range analysis suggesting that, as for the other measures, the magnitude of the effect of range decreases across sessions. Analysis of reaction-time data is pending, but the necessary data transformations of the performance records have been completed. Also planned is a retest of five of the six subjects from this experiment; the retest will occur after a delay of approximately 22 months since the end of the initial training phase of the experiment.

Analyses of the accuracy and latency measures for the Topgun part/whole training experiment have been completed. In this experiment eleven subjects were trained on both the search and the sighting components of a target engagement task on Topgun tank simulators for four sessions. Eleven other subjects were trained on only the search component of the task for the first two sessions and on both components of the task for the last two sessions. These two training groups are designated the "whole" and "part" training groups, respectively. Remarkably, part training had no detrimental effect on successful skill acquisition when measured by number of

targets destroyed. In fact, part trainers performed better than whole trainers on the first two sessions. Two measures of response latency were examined for successful kills. The first measure reflected the search component of the task (time to make an identification of the target tank), and the second measure reflected the sighting and firing component of the task (time to fire after an identification had been made). The part training group, whose search component was conducted automatically by the simulated commander for the first two sessions, performed just as well on the identification measure during the last two sessions as the whole training group, who had practice on the search component for all four sessions. In addition, though initially performing slightly worse than the whole training group on the time to fire measure in Session 1, the part training group maintained a consistent advantage over the whole training group on this measure throughout the last three sessions. Some training convergence, however, is suggested in the retention session. Latency measures on the whole suggest that part training is at least as effective as whole training in this situation when the part task is defined as a realistic subcomponent of the whole task that is neither too easy nor too difficult.

Among a college population gender differences were evident with males acquiring and destroying targets more reliably than females.

To facilitate comparisons of task-condition (part or whole training) performance in the part/whole experiment, several commonly-used methods of task analysis were explored (see Stammers & Patrick, 1975). The data may suggest that the effectiveness of part-training procedures is related to the complexity of the whole task, a hypothesis first asserted by Naylor and Briggs in 1963.

Bill Marmie is the primary investigator on this project, with Alice Healy as the faculty advisor.

Morse code reception. We completed the statistical analyses on accuracy for our follow-up study assessing the effects of confusability on retention. A set of codes, half of which are easily confused with other members of the set and half of which are not, were taught to subjects. There were two conditions: In the first condition, participants spent the acquisition period on the initial day practicing only the less-confusable letters; in the second condition, participants practiced only the more-confusable letters on the first day. On the second, third, and retention days, all participants practiced all letters. Of most interest, the accuracy results do not parallel the findings of Pellegrino, Doane, Fischer, and Alderton's (Journal of Experimental Psychology: Human Perception and Performance, 1991) visual discrimination study. The main effect of training condition was not significant during acquisition. Planned comparisons further showed that the effect of training condition was not significant at the end of training, nor was it significant in predicting the amount of forgetting during the retention period (although each group did show a temporary advantage for their initially-trained set directly after the initial session). At retention there was an advantage for the low-confusability group on the low-confusability subset, with no similar advantage for the

high-confusability group on the high-confusability subset. Throughout training, the effect of letter confusability also remained significant, with the less confusable letters showing an overall advantage. Supplementary analysis, using arcsin transformations, of the accuracy data yielded results parallel to those found on untransformed proportions. Analysis of reaction times for this study suggest similar effects.

Analyses of our study examining different aspects of the Morse code reception task suggest that discerning what auditory stimulus has been heard differs in many ways from the task of remembering the code-letter correspondence and also from the complete task of typing the letters corresponding to the given codes. The first difference is one of individual performance; only subjects in the group that heard codes and typed keys corresponding to the "dih" and "dah" pattern showed consistent accuracy performance throughout training and retention. That is, all subjects showed some improvement with training, but for those in the code-to-dida group, accuracy on the posttest and retention test was predictable from pretest scores (pretest and posttest scores correlated with $r^2=.89$; pretest and retention test scores correlated with $r^2=.74$). For the other tasks (dida-to-letter and code-to-letter) these correlations were nonsignificant. A second difference is in types of errors. On all three tests the errors made by subjects in the code-to-dida group were virtually all of the same length (number of elements) as the correct answer. Subjects in the other two groups made same-length errors at a rate slightly closer to chance. Within the same-length errors the proportions of 1-different, 2-different, and opposite errors give mixed support to Shepard's (1963) interpretation of data on discrimination and novice reception of Morse codes. In an effort to confirm the division of the Morse reception task into a code segmenting subtask and a letter identification subtask, we confirmed that average accuracy on the whole task (code-to-letter) can be predicted by average accuracies on the two partial tasks (code-to-dida and dida-to-letter). A further analysis is being performed to determine whether subtask errors predict whole task errors. Finally, the relative difficulty of difficult and easy code-letter pairs for each of the task groups is being examined to extract the sources of difficulty in the subtasks and to explore the principles that govern item difficulty.

Deborah Clawson is working on this project as part of her research for the Masters degree. She successfully defended her Masters thesis in April; a copy of her Masters thesis is attached here (see Appendix A). Her primary faculty advisor is Alice Healy, with Anders Ericsson and Lyle Bourne as her secondary faculty advisors.

Color naming. We have begun a new project on the training of color naming under both conditions of interference and noninterference. There is a large literature on the color-word interference task, originally studied by Stroop (1935). A comprehensive recent literature review (MacCleod, 1991) stimulated our interest in this topic because of the suggestion that training studies could resolve some of the crucial theoretical issues

surrounding this effect. The Stroop task also provides a well-established measure of automaticity, so that this investigation will allow us to collect important evidence on the relation between automaticity and long-term retention, which has been one of the long-standing goals of our research program.

On the basis of an extensive literature review by a large number of members of our research group, we designed and conducted an intensive training study which involved different groups of subjects exposed to two different acquisition conditions. One group of subjects was trained on naming color patches, whereas a second group of subjects was trained on naming the color of incongruous color words (i.e., the classic Stroop task). A third group of subjects was a control, receiving no training. An on-line computer program was employed in order to facilitate trial-by-trial and reaction-time analyses of the data. Twelve sessions of training were administered to each subject. Each session included 10 blocks of 24 trials, with 8 presentations of each of 3 colors in each block. All subjects, including the control subjects, were exposed to a pretest before the start of training, a posttest at the end of training, and a retention test after a one-month delay interval, which consisted of the following parts: (a) A standard paper version of the Stroop task, (b) naming of color patches, (c) color naming of incongruous color words, (c) reading of white-on-black color words, noncolor words, and nonwords, (d) reading of incongruous color words, and (e) color naming of incongruous color words under unusual orthographic manipulations designed to degrade word shape (e.g., with asterisks surrounding each letter). A standard color-vision test was administered at the start of the first session.

Initially four pilot subjects were run in our experiment. The pilot experiment gave subjects one session of practice rather than the twelve sessions planned for the actual running of the primary experiment. One session of practice does appear to affect performance in predictable ways. Performance on the interference test was improved by practice both under interference conditions and under noninterference conditions, with a speed-up of approximately 100 msec per item between the pretest and the posttest. The primary experiment including six subjects tested over either 3 or 15 sessions, is near completion. Data have been collected for all but 2 sessions for one subject, and we have begun preliminary analyses.

This research is being conducted in a collaboration by Deborah Clawson and Cheri King, with Anders Ericsson and Alice Healy as primary faculty advisors and Lyle Bourne as a secondary advisor.

Instrument panel scanning. Early this year we completed the first two of a series of experiments on the effects of random versus blocked practice on acquisition, retention, and transfer of skill, using a task that simulates an airplane pilot's job of scanning the instrument panel for emergency indications. In these two experiments we had four panels, each of which had a different rule (AND, OR, NAND, NOR) for identification of an emergency. In the first experiment one group of subjects practiced the task in a blocked format, that is, 32 trials on one rule before moving to the next

rule, whereas the other group practiced the rules in a random order. All subjects were then tested on both random and blocked trials. The subjects in the random practice group had significantly slower reaction times in the acquisition phase of the experiment, but on the two test blocks they were somewhat faster than were subjects in the blocked practice group, especially on the random test. Half of the trials included a precue (1000 ms) indicating which rule would be applicable on that trial and half had no precue (0 ms). Subjects responded faster on the trials with precues and this advantage was greater for the random group than for the blocked group.

The second experiment included a third practice group. This group was a serial group for which the rules were presented in a fixed order. That is, the first trial was in the first panel (AND), the second trial was in the second panel (OR), the third in the third panel (NAND), and the fourth in the fourth panel (NOR). Thus, subjects in this condition always knew which rule would be next. It was expected that this group would perform similarly to the random group. This experiment also included a third precue length (500 ms). The purpose of this condition was to determine if the precue needs to be long enough to allow the subject to reinstate the rule into working memory or just long enough to allow time to attend to the appropriate panel. The test for this experiment was a 144-trial random test, which was also repeated after two retention intervals. Half of the subjects returned after one week and then again three weeks later, whereas the other half returned only after one month.

The analyses show interesting results. For the acquisition and test phases of this experiment the results of the first experiment were replicated in that the blocked practice group was faster and more accurate than the random practice group during acquisition, and the trend was in the other direction for the test. The serial group, however, was halfway between the other two groups on both of the two phases. No significant differences were found for the three groups on the retention tests, although the blocked group remained slower than the other two groups. Interestingly, the blocked group's latency decreased slightly across the delay intervals, and no significant effect of forgetting was found for any of the retention intervals. The fact that the blocked group was not significantly slower than the random group on the retention test and the fact that the blocked group improved across the retention interval seems to be due to learning by the subjects on the random test immediately following the acquisition phase. This explanation is in accord with the finding by other researchers (e.g., Carlson & Yaure, 1990) that subjects first given blocked practice and then random practice are able to improve their performance over subjects given only blocked practice.

The effect of precue also replicated the first experiment in that responses were faster (and more accurate) for trials with a precue. This effect was stronger for the trials with the 1000 ms precue than for the 500 ms precue, and the effect of precue was also stronger for the random practice group as in the first experiment.

Vivian Schneider successfully completed her Dissertation Orals

and received her doctorate degree. A copy of the dissertation, which describes this study, is attached here (see Appendix B).

Later this year, as a Postdoctoral Research Associate, Dr. Schneider conducted a follow-up experiment on the effects of random versus blocked practice on acquisition, retention, and transfer of skill, using a task that simulates an airplane pilot's job of scanning the instrument panel for emergency indications. This new experiment was similar to the last one with a few changes. In this experiment subjects were trained in only blocked and random practice. Also no precues were included in this experiment; that is, there were no indications of which panel would be relevant before the trial occurred. Also, because in the last experiment subjects in the blocked practice condition improved from the immediate test to the retention test, seeming to indicate learning on the immediate random test, a new manipulation was included. Specifically, half of the subjects in both the blocked and the random practice conditions were given the immediate random test, whereas the other half were given no immediate test. A retention test was given to all subjects after a one-week delay.

The results are generally in accord with the results of the previous experiments. Subjects in the blocked practice condition were significantly faster on the acquisition phase, but were significantly slower on both the immediate and the retention test. Interestingly, on the one-week retention test, there were no significant differences between those subjects who had an immediate test and those that had no immediate test. Thus, there was no evidence of learning on the immediate test. Also, for those subjects who had both the immediate and the one-week retention test, there was no difference on the two tests, showing no forgetting.

Alice Healy is the primary faculty advisor on this research, with Anders Ericsson and Lyle Bourne as secondary advisors.

Tests of the procedural reinstatement framework. A chapter summarizing our earlier work on this topic is currently in press in a volume entitled, From Learning Processes to Cognitive Processes: Essays in Honor of William K. Estes, Volume 2. This is one of two volumes honoring William Estes and edited by Alice Healy, Stephen Kosslyn, and Richard Shiffrin. The title of the chapter is "The long-term retention of skills," and the authors are Alice Healy, David Fendrich, Robert Crutcher, William Wittman, Antoinette Gesi, Anders Ericsson, and Lyle Bourne. A copy of the chapter is attached here (see Appendix C).

A report of two subsequent experiments on this topic has been submitted for publication in a major cognitive journal. The report, which is authored by Danielle McNamara and Alice Healy, is entitled, "A generation advantage for multiplication skill and nonword vocabulary acquisition." A copy of this manuscript is attached here (see Appendix D). These two experiments were also reported by Danielle McNamara and Alice Healy at the annual meeting of the Psychonomic Society in San Francisco in November.

The first experiment in this series examined the generation effect with both simple and difficult multiplication problems. In this study 64 subjects were trained in either the read or generate condition to perform both simple and difficult multiplication problems. Subjects trained in the generate condition made significantly fewer errors on the difficult problems at the posttest than did subjects trained in the read condition. Type of training did not make a difference for accuracy performance on the simple multiplication problems. Therefore, as expected, a generation effect did occur for difficult multiplication problems, and not for simple multiplication problems.

The second study examined the long-term memorial consequences of generating versus reading nonwords. A total of 24 subjects were presented English word-nonword pairs and told that the nonwords had the same meanings in a foreign language as the corresponding English words. Subjects were trained over two consecutive days for approximately one hour each day. Subjects' recall of the nonwords was tested on the first day before training (after 10 minutes exposure to the word-nonword pairs), on the second day after completing training, and again one week following the training sessions. Those subjects who had generated the nonwords during training had significantly better long-term retention of the nonwords than did those who had simply copied the pairs. In addition, it was found that within both the read and generate conditions, those subjects who had spontaneously utilized mnemonic strategies to learn and retrieve the nonwords had significantly better recall than did those who had not. In another analysis, we addressed the question of whether the likelihood of recalling a particular nonword depended upon whether or not subjects employed a mnemonic code for that item. We found that, indeed, the proportion of correct responses was highest for items for which subjects reported using a semantic mnemonic code to remember the item, next highest for items for which subjects used a phonemic or graphemic mnemonic code, and lowest for items for which no mnemonic code was reported to be used by subjects. Also, interestingly, we found that forgetting across the one-week retention interval was smaller for items with a semantic mnemonic code than for other items, suggesting that a mnemonic coding strategy leads to particularly durable long-term retention.

The findings of both studies described above are understood in terms of theories of proceduralization. It is hypothesized that subjects in a generate condition are more likely to develop cognitive procedures during training than are subjects in a read condition. These procedures, developed during training, are subsequently reinstated at test. We believe that the significant factor in the generation effect is that subjects develop cognitive procedures to learn and recall the test items. Both of the studies described above support this hypothesis.

A follow-up experiment was completed which further investigated the generation effect with simple and difficult multiplication problems. Subjects read aloud the problems and either read or generated the answers. After a brief interpolated task, their memory for the answers was tested, first with a recall procedure and then

with a recognition procedure. Our procedural account led to the prediction of a larger generation effect for simple than for difficult multiplication problems in this task which tests episodic memory rather than skill acquisition because the subjects would be likely to have stronger memorial links between the operands and the products for the simple than for the difficult problems. However, our findings showed a generation effect for both types of problems. This experiment was the basis for a senior undergraduate honors thesis by David McCain. His advisors were Alice Healy and Danielle McNamara, who are also conducting a second follow-up experiment with the aim to throw additional light on the unexpected finding of a generation effect on episodic memory for the answers to difficult multiplication problems. The new experiment includes addition problems as well as multiplication problems.

A related study on the generation effect was completed in collaboration with J. David Mason, a graduate student in the Accounting Department of the School of Business at the University of Colorado. This study included two experiments in which college students were given accounting problems with simple and complex numbers. Even when the to-be-remembered information was equated for both conditions, memory was greater for simple than for complex problems by both recall and recognition measures. A manuscript describing this research was submitted for publication by David Mason, Alice Healy, and Bill Marmie. A copy of the manuscript is attached here (see Appendix E).

Danielle McNamara is currently the primary investigator on this research; Alice Healy is the primary faculty advisor.

Mental calculation. Some of our earlier work on this topic was summarized in a paper by Lyle Bourne and Tim Rickard delivered at the Inter-American Congress of Psychology, San Jose, Costa Rica.

Tim Rickard successfully defended his Masters thesis this semester. The thesis reported two of our earlier studies on the transfer of arithmetic skill. A copy of the thesis is attached here (see Appendix F).

Data from three new mental arithmetic studies will be available by early May. Two of these are extensions of our previous explorations of practice and transfer of arithmetic skill. Subjects will be trained on a subset of simple multiplication and division problems, and then tested on the exact problems on which they practiced, as well as on various other versions of the practice problems. One group of subjects will be tested immediately after practice and after a one-month retention interval, and another group will be tested only after the one-month delay. The overall goals of these practice/transfer studies are (a) to understand in some detail the nature of the representation of arithmetic facts in memory, and (b) to explore further retention of any improvements in performance gained with practice. With respect to retention, our previous research showed excellent retention of skill for arithmetic problems on which subjects practiced. In the current study, we will attempt to replicate this finding, and also to determine whether learning

which transfers to the altered problems is also retained over the long term.

The third study explores whether the problem size effect in mental multiplication (the fact that small problems such as 3×4 are solved more accurately and quickly than large problems such as 6×9) disappears after extensive practice. The character and extent of change in the problem size effect with practice should yield insight into the cognitive mechanisms that underlie this robust effect. Contingent on the results, subjects may be called back after a delay to be retested. Based on our previous work, we expect that any decrement in the problem-size effect which occurs with practice will be localized in the calculation (or retrieval) component of performance (not the motor programming component) and will be retained over the retention interval.

An additional arithmetic project which we are currently pursuing is the development of a connectionist model of skilled performance in multiplication. The model is based on the premise that skilled arithmetic involves retrieval of facts from an associative network, with occasional use of back-up strategies when retrieval fails. Simulations to date show that the model should be able to account for the various reaction time and error phenomena which have been documented for skilled multiplication. We have completed a preliminary manuscript of this work, and are currently expanding on the existing simulations.

As a follow-up to our previous research on arithmetic, we designed and conducted an experiment which explored strategies that college students use to solve multiplication problems. Recently, researchers have argued that college students are at a skill level that allows them to retrieve answers to simple multiplication problems (e.g., $7 \times 8 = _$) directly from an associative network. To investigate this hypothesis, we had college subjects solve multiplication problems, and then tell us after each problem whether they spontaneously knew the answer on seeing the problem (this response would indicate network retrieval), or had to do some thinking before coming up with the answer. Results showed that many subjects did not spontaneously know the answer to some problems. For typically more difficult problems (e.g., 6×9), this finding was true nearly half the time across subjects. Also, the data showed that subjects did not spontaneously arrive at the answer to some problems even after 30 exposures. Preliminary analyses suggest that some arithmetic phenomena that have previously been attributed to the retrieval process may actually be a consequence of the strategies that are sometimes used to generate the answer. For example, the well established problem-size effect seems to be largely, if not completely, a result of non-retrieval processes. In follow-up research currently underway, we are attempting to identify specific non-retrieval strategies that are used, and to establish the extent to which these strategies are responsible for performance phenomena such as the problem-size effect. Long-term retention tests are included in these studies so that we can evaluate the degree of consistency of strategy use across time. The first study along these lines involves collection of protocols of individual strategy use on

a problem by problem basis. Preliminary results suggest substantial variation of strategy use across both subjects and problems.

Tim Rickard is the primary investigator of this research. His primary faculty advisor is Lyle Bourne, with Alice Healy and Anders Ericsson as secondary advisors.

Memory for instances of categories. We designed and conducted a set of studies to explore the acquisition and retention of conceptual knowledge for faces. In the original pilot study, subjects were shown a set of faces that had similar features (i.e., family faces), followed by a set of faces that were dissimilar from the faces in the first set and from each other (i.e., non-family faces). Subjects were then tested on selected faces from both of these sets, as well as on new family faces not previously seen. During the test, subjects were asked to identify faces that were old family faces. Preliminary results demonstrate that subjects are slightly more likely to identify family faces (both old and new) than non-family faces at test. There was no evidence for a tendency to identify old non-family faces more than new non-family faces. Results indicate that memory for the faces was strongly feature-driven, and that the particular faces used in each phase of the experiment had a major impact on the results.

In follow-up research, we are experimenting with using non-overlapping feature pools for family and non-family faces in an effort to get more substantial differences across conditions. We are also manipulating the salience level of the faces in each group (i.e., how good of an exemplar of its group is a given face) in order to test the notion of feature-driven category representation. We will also test some subjects immediately after acquisition, and others after a one-day or two-week delay. We expect to find evidence that, on the immediate test, subjects will be able to discriminate family from non-family faces, but that after a delay, subjects will have lost the memory for context (i.e., family or non-family) and will base their responses on familiarity only. Thus, we expect that subjects will mistakenly identify non-family faces as family faces more often at delay than at immediate test. At this point, the data have been collected and the analyses are underway.

This study is the basis of a first-year research project by Oliver Hammerle. The results of this study will be the subject of an upcoming talk by Oliver Hammerle at the University of Colorado Psychology Department's 11th Annual Mini-Convention.

Oliver Hammerle and Tim Rickard are engaged in this research with Lyle Bourne as their primary faculty advisor.

Target detection. The article by Vivian Schneider, Alice Healy, and Antoinette Gesi summarizing some of our earlier experiments on target detection was published this year in the Journal of Memory and Language.

This year we also revised a manuscript summarizing some of our earlier studies on letter detection conducted in collaboration with

Jeffrey Hadley, who is currently a postdoctoral researcher at the National Institutes of Health. This manuscript by Jeffrey Hadley and Alice Healy was accepted for publication and published this year in the Journal of Experimental Psychology: Learning, Memory, and Cognition.

In addition, this year we revised a manuscript summarizing some of our earlier studies on proofreading conducted in collaboration with Thomas Cunningham, who is a Professor at St. Lawrence University. This manuscript by Alice Healy and Thomas Cunningham was accepted for publication and recently published in the journal Memory & Cognition. A copy of this article is attached here (see Appendix G).

Alice Healy completed a major paper reviewing our research on target detection. This paper was to be delivered as an invited address at the American Psychological Association convention, which was held in San Francisco in August. However, because of serious medical problems, Professor Healy was not able to deliver her address. As mentioned above, she did deliver the paper at the Institute of Cognitive Science in September. Also, she plans to revise the paper so that it will be suitable for publication as a book chapter.

Bill Oliver, who is now an Assistant Professor at Florida State University, visited in December and worked on a revision with Alice Healy of a manuscript of theirs summarizing earlier work on the letter detection task.

Vivian Schneider and Alice Healy completed a new experiment in their continuing series on phonetic factors influencing letter detection. In this experiment the importance of the location of the target letter was explored by comparing letter detection errors for the targets t and o in the the words to, two, and not. The location of the target letter in the word was found to be less important than the frequency of the word and the nature of the target letter. Specifically, subjects made more errors on the more common word to than on the less common word two, and subjects made more errors when searching for the target t than when searching for the target o. Our working hypothesis about the latter finding is that the special visual configurational properties of the letter o make it particularly easy to detect.

To test this hypothesis concerning the visual nature of the target letter, Vivian Schneider and Alice Healy conducted a second new experiment with two different versions of the passage used in the earlier experiment. Half the subjects received one version of the passage, and the other half of the subjects received the other version of the passage. One version was typed in an ordinary font (Times with New York spacing), with the letter o perfectly symmetrical, and the other version was typed in an unusual font (Venice), with a somewhat asymmetrical o. We found that detection error rates did increase with the unusual font relative to the standard font, but the increase was caused by more errors with the target t as well as with the target o. This result can be explained

by the fact that the visual form of the letter t as well as that of the letter o is less regular in the unusual font.

A follow-up experiment was conducted which was identical to the recently completed experiment except that a new font (Nordic) was employed which has asymmetrical os but regular ts. We found, as expected, that the detection error rate was increased for os but decreased for ts with this new font, thereby demonstrating the important role of type font in letter detection.

Vivian Schneider and Alice Healy also conducted another experiment with a similar manipulation of type font to explore whether the unitization of a word would be affected by its visual configuration. Specifically, we compared letter detection in the highly frequent word the to less frequent words using passages typed normally (in the Times font with New York spacing) or typed in script font (Florence). We found detection errors decreased in the script font for both frequent and rare test words. This finding is consistent with the argument that the familiarity of the visual configuration of the word leads the word to be treated as a unit, thus making it more difficult to detect individual letters in the word. This finding is particularly noteworthy because although it is consistent with the unitization theoretical framework it is inconsistent with intuition. Subjects found it easier to detect a letter when it was typed in an unusual fashion (i.e., script) presumably because not only was the letter unusual but so was the larger word unit, so that unitization was less likely.

Alice Healy is the primary faculty investigator for this line of research.

Data entry. We worked on revising our manuscript with Antoinette Gesi, who is now a graduate student at the University of California, Santa Cruz, and David Fendrich, who is now an Assistant Professor at Widener University, summarizing one of our studies on data entry. Both Ms. Gesi and Dr. Fendrich visited us this year to facilitate our progress in revising this manuscript. In our visits we reached the conclusion that a follow-up experiment was needed for this study. We designed the follow-up experiment, and Dr. Fendrich began conducting that experiment this year at Widener University. The primary faculty advisor on this research project has been Alice Healy, with Lyle Bourne as the secondary advisor.

Temporal, spatial, and item components of memory for lists. Thomas Cunningham presented a report at the November meeting of the Psychonomic Society summarizing some of our earlier research on this topic. The authors of the report are Thomas Cunningham, Alice Healy, Robert Till, David Fendrich, and Christina Dimitry. A second report on previous research on this topic was presented at the Psychonomic Society meetings by Robert Till of the University of North Dakota. The authors of that report are Robert Till, Alice Healy, Thomas Cunningham, and Lyle Bourne. A third report of earlier research on this topic was published this year in a chapter by Alice Healy, Thomas Cunningham, Antoinette Gesi, Robert Till, and Lyle Bourne in a volume honoring Bennet B. Murdock.

This year we also worked on a related paper concerning short-term memory. This paper discusses research conducted in collaboration with Jeffrey Hadley and Bennet B. Murdock, a Professor at the University of Toronto. This paper by Jeffrey Hadley, Alice Healy, and Bennet Murdock was revised subsequent to a very favorable editorial review and was recently published in Memory & Cognition.

An experiment has been conducted which examined the impact on short-term recall of order information created by manipulating the physical attributes of a single item presented in a sequence of taxonomically similar items, that is, letters of the alphabet. Alphanumeric characters were presented successively on a CRT screen for brief periods of time. An experimental trial included the presentation of a pair of four-letter segments followed by a digit-filled retention interval and cue to indicate that either the first or second four-letter segment should be recalled. This recall cue was presented to the subject both at the start and at the end of the trial. A single letter in the to-be-recalled segment was made distinctive by increasing its brightness level and introducing a unique hue. Similar items were in black and distinct items were in red. Six variables were manipulated: (a) exposure time -- 500 ms or 1000 ms, (b) instructions to subjects -- emphasis or no emphasis on importance of recall of distinctive letter, (c) letter distinctiveness -- present or absent, (d) retention interval -- 0, 4, or 12 interpolated digits, (e) segment cued for recall -- first or second, and (f) letter position -- first, second, third or fourth.

Thirty-two subjects were tested in this experiment. Subjects were instructed to recall one of two four-item segments of letters. On half of the 96 experimental trials one item in the to-be-recalled segment was displayed in bright red (as opposed to black). Analyses of the effects of retention interval, segment cued for recall, serial position, and presentation rate were generally consistent with previous findings in our laboratory (see, e.g., Cunningham, Healy, & Williams, 1984; Healy, Fendrich, Cunningham, & Till, 1987). Analyses also revealed that no overall enhancement of segment retention was evident for the distinctive trials although segment cued for recall significantly interacted with letter distinctiveness so that for the first segment recall was better when the trial contained no distinct item but for the second segment there was a trend in the opposite direction. Further analyses examined whether subjects were able to position distinctive items better, whether an item's position interacted with its distinctiveness, and whether an overall elevation of retention for a distinctive item was achieved at the cost of an overall depression of retention for the non-distinct items in a given segment. These analyses, which were conducted only on trials including a distinct letter, revealed a three-way interaction of segment cued for recall, serial position, and the distinct/nondistinct status of a given letter. There was an advantage in recall level for a distinct letter if it was in the first position of the first segment or in the middle positions of the second segment. We also found a three-way interaction of retention interval, serial position, and the distinct/nondistinct status of a given letter. We found that when there were no intervening digits, there was an advantage for the distinct letter in the first serial

position, whereas when there were 12 intervening digits, there was an advantage for the distinct letter in the last serial position. We plan to explore variations of the Estes perturbation model of short-term ordered recall in an attempt to provide an account for this complex pattern of results. The variation of this model we have been using contains both short-term and long-term memory components, so that we should be able to determine whether the effects of distinctiveness are attributable largely to short- or long-term memorial processes. The results of this experiment will be the subject of an upcoming talk by Bill Marmie at the University of Colorado Psychology Department's 11th Annual Mini-Convention.

A follow-up experiment will be conducted using a Macintosh computer system. From this study we hope to replicate the previous experiment and examine the order in which subjects recall each letter of a segment. In particular, we will be interested in determining whether subjects recall the distinctive item before the non-distinct items. A future study will then look at the effect of directing a subject to recall the distinctive item first.

We prepared the stimulus materials for a follow-up experiment which we sent to our colleague Thomas Cunningham at St. Lawrence University who is conducting the testing of subjects there. In our recently completed experiment the letter was shown, whereas in the follow-up experiment it is being replaced by a single dash. In both cases the distinctive brightness and hue are utilized. Although the distinctive letter is missing from the follow-up experiment, it can be generated by the subject who is familiar with the constrained letter set employed in the experiment.

Bill Marmie is working on this project as the basis for both his first-year research project and his Master's thesis. Alice Healy is his principal faculty advisor and Tom Cunningham is his secondary advisor.

A new line of research was begun which directly compares short- and long-term retention of features of common objects. The figures and words of two United States coins, a modern penny, and a Liberty dime, were used as stimuli. A pilot study, using 8 subjects, was conducted in the Fall of 1991. A full experiment, using 32 subjects, was recently completed. In both studies each subject drew both a penny and a Liberty dime. The penny was drawn entirely from memory with no visual cues either given or allowed. The dime was drawn after one minute of study time was allowed. One manipulation of interest in the full experiment was the retention interval between study of the dime and recall of its features. One group of subjects drew the dime immediately after studying it (the short-term retention group), and another group of subjects drew it one week later (the long-term retention group). Results from this experiment indicated that subjects showed better memory for features of the Liberty dime than features of the penny even when the dime was recalled one week after studying it, and even though the penny had received presumably much greater than a single minute of study during the everyday life of the subject. This finding points to important differences between semantic memory (responsible for the retention of the penny) and

episodic memory (responsible for the retention of the Liberty dime).

The inclusion of the penny in the experiment extends the research of Nickerson and Adams (1979) who demonstrated poorer recall for the features of a common U. S. penny than one may have expected from an object seen (presumably) quite often. By counterbalancing presentation and study of the dime prior to or following recall of the penny, we are exploring the degree to which prior information of featural elements transfers from knowledge of the dime to knowledge of the penny. We have found that this manipulation reveals a reliable advantage for the penny when it follows the dime relative to when it precedes the dime. Conclusions regarding the short- and long-term retention comparison as well as extensions of Nickerson and Adams' (1979) theoretical accounts will be forthcoming.

This project is being conducted as an independent project by Bill Marmie and Gregory Rully with Alice Healy as the primary advisor.

Grant Sinclair and Alice Healy have been working on a manuscript summarizing our earlier research conducted in collaboration with Lyle Bourne comparing the learning and long-term retention of temporal, spatial, and item information.

Temporal, spatial, and item components of memory for course schedules. This year Major William Wittman, who is currently on the faculty of the Air Force Academy, and Alice Healy continued their work on a manuscript summarizing our earlier research on this topic. This manuscript is being prepared for submission to a cognitive journal.

This year we also conducted a study on this topic in collaboration with Cheri King, who is a doctoral student at Colorado State University and is performing this research as her doctoral dissertation. In this study, students from both the University of Colorado and Colorado State University participated as subjects and learned course schedules from either their own institution or from the unfamiliar campus. Each of the subjects learned a course schedule in one session, followed by retention tests at one-week and either two-, three-, or four-week intervals. Ninety-six subjects were tested. In accord with our procedural reinstatement framework, we expected superior memory for spatial information about course locations only for subjects familiar with the relevant campus. Progress was made in analyzing the data from the experiment. There was a significant loss in retention from test to retest. On the class listing tests, subjects in the familiar condition had better recall of spatial information than subjects in the unfamiliar condition. However, there was no effect of superior memory for spatial information over item and temporal information.

Alice Healy is the primary faculty advisor for this project, and Lyle Bourne is a secondary advisor.

Vocabulary retention. During the past year, analyses and retention testing were completed for an experiment involving over 50

blocks of practice on Spanish vocabulary items, which had been learned with the keyword method. Our results show that before training subjects retrieved the English words by sequentially generating the keyword and then the English word. This finding was supported by analyses of both response time (RT) and retrospective verbal reports. After extended training subjects retrieved the English word directly, as shown by faster access to the English word when cued with the Spanish word than when cued with the keyword. Retrospective reports on retrieval after training confirm the direct access after training.

Our previous analyses showed that before extended retrieval practice, subjects retrieve the English equivalents of Spanish-English pairs learned using the keyword method by first retrieving the keyword from the Spanish word and then using the keyword to retrieve the English equivalent. However, following extended retrieval practice, subjects are able to access directly the English equivalents from the Spanish word, as evidenced by: (a) faster access to the English word when cued with the Spanish word than when cued with the keyword; and (b) a dramatic decline in verbal reports of keyword mediators. Recently we analyzed the one-month retention data for this experiment. These analyses were consistent with those for the immediate results, supporting a direct access account. Whereas cued recall was at ceiling for the immediate results, after the one-month delay, recall declined enough to compare performance on the different cued recall tasks. Cued recall for the English equivalent was higher given the Spanish word as a retrieval cue than given the keyword as a cue, consistent with the direct access account. In addition, access to the English word via the Spanish word was still faster than access via the keyword, although only marginally.

We also conducted and analyzed the results for two new experiments exploring the role of mediation in retrieval of foreign vocabulary. These studies employed an interference paradigm to examine the continued involvement of keyword mediators in the retrieval process. The paradigm involves learning new associates to the earlier-used keywords for learning the Spanish-English vocabulary pairs. The first experiment examined such interference effects prior to extended retrieval practice when retrieval of the English equivalents is still mediated by first retrieving the keyword. The second experiment examined these interference effects after extended retrieval practice of the Spanish-English pairs when retrieval of the English equivalents is direct.

Analyses of the results for the first experiment showed a considerable interference effect. Retrieval of the English equivalent was considerably slowed by having learned a new associate to the previously-learned keyword--as would be expected if retrieval of the English equivalent depends on accessing the keyword. However, the results for the second experiment were quite surprising. As in the earlier extended practice experiment, following extended retrieval practice, subjects were faster retrieving the English equivalent from the Spanish word than from the keyword and no longer reported keyword mediators in accessing the English definitions.

Nevertheless, learning new associates to the old keywords reliably slowed retrieval of the English equivalents of the Spanish words, suggesting that the keywords may still be involved in the retrieval process, even though subjects are no longer consciously aware of retrieving them.

During the coming year, we plan to do further analyses on these data and to begin writing up the results of these and earlier experiments for publication.

The work on this topic was the basis for a doctoral dissertation by Robert Crutcher. Robert Crutcher successfully defended his dissertation this year; a copy of the dissertation is enclosed here (see Appendix H). His primary faculty advisor was Anders Ericsson; Alice Healy was a secondary advisor. Robert Crutcher also gave two colloquia describing this research, one at Denver University and the other at the University of Illinois at Chicago, where Dr. Crutcher will be an Assistant Professor starting this Fall.